



#5

Sequence Listing

(1) GENERAL INFORMATION:

(i) APPLICANT: The Government of the United States of America
as represented by the Secretary
Department of Health and Human Services
Washington, D.C.

Htun Ph.D., Han
Hager Ph.D., Gordon L.

(ii) TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MONITORING
DNA BINDING MOLECULES IN LIVING CELLS

(iii) NUMBER OF SEQUENCES: 9

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Needle & Rosenberg
(B) STREET: 127 Peachtree Street, Suite 1200
(C) CITY: Atlanta
(D) STATE: Georgia
(E) COUNTRY: USA
(F) ZIP: 30303

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 60/008,373
(B) FILING DATE: 08 Dec 1995
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Selby, Elizabeth
(B) REGISTRATION NUMBER: 38298
(C) REFERENCE/DOCKET NUMBER: 14014.0183

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 404-688-0770
(B) TELEFAX: 404-688-9880

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 7257 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: double
(D) Topology: linear

(ii) MOLECULE TYPE: nucleic acid

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 1072..4284

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TCAATATTGG CCATTAGCCA TATTATTCAT TGGTTATATA GCATAAATCA ATATTGGCTA	60
TTGGCCATTG CATACGTTGT ATCTATATCA TAATATGTAC ATTTATATTG GCTCATGTCC	120
AATATGACCG CCATGTTGGC ATTGATTATT GACTAGTTAT TAATAGTAAT CAATTACGGG	180
GTCATTAGTT CATAGCCCAT ATATGGAGTT CCGCGTTACA TAACTTACGG TAAATGGCCC	240
GCCTGGCTGA CCGCCCAACG ACCCCCGCCC ATTGACGTCA ATAATGACGT ATGTTCCCAT	300
AGTAACGCCA ATAGGGACTT TCCATTGACG TCAATGGGTG GAGTATTTAC GGTAAACTGC	360
CCACTTGGCA GTACATCAAG TGTATCATAT GCCAAGTCCG CCCCTATTG ACGTCAATGA	420
CGGTAAATGG CCCGCTGGC ATTATGCCCC GTACATGACC TTACGGGACT TTCCTACTTG	480
GCAGTACATC TACGTATTAG TCATCGCTAT TACCATGGTG ATGCGGTTTT GGCAGTACAC	540
CAATGGGCGT GGATAGCGGT TTGACTCACG GGGATTTCCA AGTCTCCACC CCATTGACGT	600
CAATGGGAGT TTGTTTTGGC ACCAAAATCA ACGGGACTTT CCAAATGTC GTAATAACCC	660
CGCCCCGTTG ACGCAAATGG GCGGTAGGCG TGTACGGTGG GAGGTCTATA TAAGCAGAGC	720
TCGTTTAGTG AACCCTCAGA TCACTAGAAG CTTTATTGCG GTAGTTTATC ACAGTTAAAT	780
TGCTAACGCA GTCAGTGCTT CTGACACAAC AGTCTCGAAC TTAAGCTGCA GAAGTTGGTC	840
GTGAGGCACT GGGCAGGTAA GTATCAAGGT TACAAGACAG GTTTAAGGAG ACCAATAGAA	900
ACTGGGCTTG TCGAGACAGA GAAGACTCTT GCGTTTCTGA TAGGCACCTA TTGGTCTTAC	960
TGACATCCAC TTGCTTTTC TCTCCACAGG TGTCCACTCC CAGTTCAATT ACAGCTCTTA	1020
AGGCTAGAGT ACTTAATACG ACTCACTATA GGCTAGCGAA GGAGATCCGC C ATG GCC	1077
	Met Ala
	1
CAC CAT CAC CAC CAT CAC GGA TAT CCA TAC GAC GTG CCA GAT TAC GCT	1125
His His His His His His Gly Tyr Pro Tyr Asp Val Pro Asp Tyr Ala	
5 10 15	
CAG TCG AGT GCC ATG AGT AAA GGA GAA GAA CTT TTC ACT GGA GTT GTC	1173
Gln Ser Ser Ala Met Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val	
20 25 30	

CCA ATT CTT GTT GAA TTA GAT GGT GAT GTT AAT GGG CAC AAA TTT TCT	1221
Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser	
35 40 45 50	
GTC AGT GGA GAG GGT GAA GGT GAT GCA ACA TAC GGA AAA CTT ACC CTT	1269
Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu	
55 60 65	
AAA TTT ATT TGC ACT ACT GGA AAA CTA CCT GTT CCT TGG CCA ACA CTT	1317
Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu	
70 75 80	
GTC ACT ACT TTC ACT TAT GGT GTT CAA TGC TTT TCA AGA TAC CCA GAT	1365
Val Thr Thr Phe Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp	
85 90 95	
CAT ATG AAA CAG CAT GAC TTT TTC AAG AGT GCC ATG CCC GAA GGT TAT	1413
His Met Lys Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr	
100 105 110	
GTA CAG GAA AGA ACT ATA TTT TTC AAA GAT GAC GGG AAC TAC AAG ACA	1461
Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr	
115 120 125 130	
CGT GCT GAA GTC AAG TTT GAA GGT GAT ACC CTT GTT AAT AGA ATC GAG	1509
Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu	
135 140 145	
TTA AAA GGT ATT GAT TTT AAA GAA GAT GGA AAC ATT CTT GGA CAC AAA	1557
Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys	
150 155 160	
TTG GAA TAC AAC TAT AAC TCA CAC AAT GTA TAC ATC ATG GCA GAC AAA	1605
Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys	
165 170 175	
CAA AAG AAT GGA ATC AAA GTT AAC TTC AAA ATT AGA CAC AAC ATT GAA	1653
Gln Lys Asn Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu	
180 185 190	
GAT GGA AGC GTT CAA CTA GCA GAC CAT TAT CAA CAA AAT ACT CCA ATT	1701
Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile	
195 200 205 210	
GGC GAT GGC CCT GTC CTT TTA CCA GAC AAC CAT TAC CTG TCC ACA CAA	1749
Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln	
215 220 225	
TCT GCC CTT TCG AAA GAT CCC AAC GAA AAG AGA GAC CAC ATG GTC CTT	1797
Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu	
230 235 240	
CTT GAG TTT GTA ACA GCT GCT GGG ATT ACA CAT GGC ATG GAT GAA CTA	1845
Leu Glu Phe Val Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu Leu	
245 250 255	

TAC AAA GGC GCC GGC GCT GGT GCT GGT GCT GGC GCC ATC AGC GCG CTG	1893
Tyr Lys Gly Ala Gly Ala Gly Ala Gly Ala Gly Ala Ile Ser Ala Leu	
260 265 270	
ATC CTG GAC TCC AAA GAA TCC TTA GCT CCC CCT GGT AGA GAC GAA GTC	1941
Ile Leu Asp Ser Lys Glu Ser Leu Ala Pro Pro Gly Arg Asp Glu Val	
275 280 285 290	
CCT GGC AGT TTG CTT GGC CAG GGG AGG GGG AGC GTA ATG GAC TTT TAT	1989
Pro Gly Ser Leu Leu Gly Gln Gly Arg Gly Ser Val Met Asp Phe Tyr	
295 300 305	
AAA AGC CTG AGG GGA GGA GCT ACA GTC AAG GTT TCT GCA TCT TCG CCC	2037
Lys Ser Leu Arg Gly Gly Ala Thr Val Lys Val Ser Ala Ser Ser Pro	
310 315 320	
TCA GTG GCT GCT GCT TCT CAG GCA GAT TCC AAG CAG CAG AGG ATT CTC	2085
Ser Val Ala Ala Ala Ser Gln Ala Asp Ser Lys Gln Gln Arg Ile Leu	
325 330 335	
CTT GAT TTC TCG AAA GGC TCC ACA AGC AAT GTG CAG CAG CGA CAG CAG	2133
Leu Asp Phe Ser Lys Gly Ser Thr Ser Asn Val Gln Gln Arg Gln Gln	
340 345 350	
CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG	2181
Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln	
355 360 365 370	
CAG CCA GGC TTA TCC AAA GCC GTT TCA CTG TCC ATG GGG CTG TAT ATG	2229
Gln Pro Gly Leu Ser Lys Ala Val Ser Leu Ser Met Gly Leu Tyr Met	
375 380 385	
GGA GAG ACA GAA ACA AAA GTG ATG GGG AAT GAC TTG GGC TAC CCA CAG	2277
Gly Glu Thr Glu Thr Lys Val Met Gly Asn Asp Leu Gly Tyr Pro Gln	
390 395 400	
CAG GGC CAA CTT GGC CTT TCC TCT GGG GAA ACA GAC TTT CGG CTT CTG	2325
Gln Gly Gln Leu Gly Leu Ser Ser Gly Glu Thr Asp Phe Arg Leu Leu	
405 410 415	
GAA GAA AGC ATT GCA AAC CTC AAT AGG TCG ACC AGC GTT CCA GAG AAC	2373
Glu Glu Ser Ile Ala Asn Leu Asn Arg Ser Thr Ser Val Pro Glu Asn	
420 425 430	
CCC AAG AGT TCA ACG TCT GCA ACT GGG TGT GCT ACC CCG ACA GAG AAG	2421
Pro Lys Ser Ser Thr Ser Ala Thr Gly Cys Ala Thr Pro Thr Glu Lys	
435 440 445 450	
GAG TTT CCC AAA ACT CAC TCG GAT GCA TCT TCA GAA CAG CAA AAT CGA	2469
Glu Phe Pro Lys Thr His Ser Asp Ala Ser Ser Glu Gln Gln Asn Arg	
455 460 465	
AAA AGC CAG ACC GGC ACC AAC GGA GGC AGT GTG AAA TTG TAT CCC ACA	2517
Lys Ser Gln Thr Gly Thr Asn Gly Gly Ser Val Lys Leu Tyr Pro Thr	

470

475

480

GAC CAA AGC ACC TTT GAC CTC TTG AAG GAT TTG GAG TTT TCC GCT GGG Asp Gln Ser Thr Phe Asp Leu Leu Lys Asp Leu Glu Phe Ser Ala Gly	2565
485 490 495	
TCC CCA AGT AAA GAC ACA AAC GAG AGT CCC TGG AGA TCA GAT CTG TTG Ser Pro Ser Lys Asp Thr Asn Glu Ser Pro Trp Arg Ser Asp Leu Leu	2613
500 505 510	
ATA GAT GAA AAC TTG CTT TCT CCT TTG GCG GGA GAA GAT GAT CCA TTC Ile Asp Glu Asn Leu Leu Ser Pro Leu Ala Gly Glu Asp Asp Pro Phe	2661
515 520 525 530	
CTT CTC GAA GGG AAC ACG AAT GAG GAT TGT AAG CCT CTT ATT TTA CCG Leu Leu Glu Gly Asn Thr Asn Glu Asp Cys Lys Pro Leu Ile Leu Pro	2709
535 540 545	
GAC ACT AAA CCT AAA ATT AAG GAT ACT GGA GAT ACA ATC TTA TCA AGT Asp Thr Lys Pro Lys Ile Lys Asp Thr Gly Asp Thr Ile Leu Ser Ser	2757
550 555 560	
CCC AGC AGT GTG GCA CTA CCC CAA GTG AAA ACA GAA AAA GAT GAT TTC Pro Ser Ser Val Ala Leu Pro Gln Val Lys Thr Glu Lys Asp Asp Phe	2805
565 570 575	
ATT GAA CTT TGC ACC CCC GGG GTA ATT AAG CAA GAG AAA CTG GGC CCA Ile Glu Leu Cys Thr Pro Gly Val Ile Lys Gln Glu Lys Leu Gly Pro	2853
580 585 590	
GTT TAT TGT CAG GCA AGC TTT TCT GGG ACA AAT ATA ATT GGT AAT AAA Val Tyr Cys Gln Ala Ser Phe Ser Gly Thr Asn Ile Ile Gly Asn Lys	2901
595 600 605 610	
ATG TCT GCC ATT TCT GTT CAT GGT GTG AGT ACC TCT GGA GGA CAG ATG Met Ser Ala Ile Ser Val His Gly Val Ser Thr Ser Gly Gly Gln Met	2949
615 620 625	
TAC CAC TAT GAC ATG AAT ACA GCA TCC CTT TCT CAG CAG CAG GAT CAG Tyr His Tyr Asp Met Asn Thr Ala Ser Leu Ser Gln Gln Gln Asp Gln	2997
630 635 640	
AAG CCT GTT TTT AAT GTC ATT CCA CCA ATT CCT GTT GGT TCT GAA AAC Lys Pro Val Phe Asn Val Ile Pro Pro Ile Pro Val Gly Ser Glu Asn	3045
645 650 655	
TGG AAT AGG TGC CAA GGC TCC GGA GAG GAC AGC CTG ACT TCC TTG GGG Trp Asn Arg Cys Gln Gly Ser Gly Glu Asp Ser Leu Thr Ser Leu Gly	3093
660 665 670	
GCT CTG AAC TTC CCA GGC CGG TCA GTG TTT TCT AAT GGG TAC TCA AGC Ala Leu Asn Phe Pro Gly Arg Ser Val Phe Ser Asn Gly Tyr Ser Ser	3141
675 680 685 690	
CCT GGA ATG AGA CCA GAT GTA AGC TCT CCT CCA TCC AGC TCG TCA GCA	3189

Pro Gly Met Arg	Pro Asp Val Ser Ser	Pro Pro Ser Ser Ser Ser Ala	
	695	700	705
GCC ACG GGA CCA CCT CCC AAG CTC TGC CTG GTG TGC TCC GAT GAA GCT			3237
Ala Thr Gly Pro Pro Pro Lys Leu Cys Leu Val Cys Ser Asp Glu Ala	710	715	720
TCA GGA TGT CAT TAC GGG GTG CTG ACA TGT GGA AGC TGC AAA GTA TTC			3285
Ser Gly Cys His Tyr Gly Val Leu Thr Cys Gly Ser Cys Lys Val Phe	725	730	735
TTT AAA AGA GCA GTG GAA GGA CAG CAC AAT TAC CTT TGT GCT GGA AGA			3333
Phe Lys Arg Ala Val Glu Gly Gln His Asn Tyr Leu Cys Ala Gly Arg	740	745	750
AAC GAT TGC ATC ATT GAT AAA ATT CGA AGG AAA AAC TGC CCA GCA TGC			3381
Asn Asp Cys Ile Ile Asp Lys Ile Arg Arg Lys Asn Cys Pro Ala Cys	755	760	765
CGC TAT CGG AAA TGT CTT CAG GCT GGA ATG AAC CTT GAA GCT CGA AAA			3429
Arg Tyr Arg Lys Cys Leu Gln Ala Gly Met Asn Leu Glu Ala Arg Lys	775	780	785
ACA AAG AAA AAA ATC AAA GGG ATT CAG CAA GCC ACT GCA GGA GTC TCA			3477
Thr Lys Lys Lys Ile Lys Gly Ile Gln Gln Ala Thr Ala Gly Val Ser	790	795	800
CAA GAC ACT TCG GAA AAT CCT AAC AAA ACA ATA GTT CCT GCA GCA TTA			3525
Gln Asp Thr Ser Glu Asn Pro Asn Lys Thr Ile Val Pro Ala Ala Leu	805	810	815
CCA CAG CTC ACC CCT ACC TTG GTG TCA CTG CTG GAG GTG ATT GAA CCC			3573
Pro Gln Leu Thr Pro Thr Leu Val Ser Leu Leu Glu Val Ile Glu Pro	820	825	830
GAG GTG TTG TAT GCA GGA TAT GAT AGC TCT GTT CCA GAT TCA GCA TGG			3621
Glu Val Leu Tyr Ala Gly Tyr Asp Ser Ser Val Pro Asp Ser Ala Trp	835	840	845
AGA ATT ATG ACC ACA CTC AAC ATG TTA GGT GGG CGT CAA GTG ATT GCA			3669
Arg Ile Met Thr Thr Leu Asn Met Leu Gly Gly Arg Gln Val Ile Ala	855	860	865
GCA GTG AAA TGG GCA AAG GCG ATA CTA GGC TTG AGA AAC TTA CAC CTC			3717
Ala Val Lys Trp Ala Lys Ala Ile Leu Gly Leu Arg Asn Leu His Leu	870	875	880
GAT GAC CAA ATG ACC CTG CTA CAG TAC TCA TGG ATG TTT CTC ATG GCA			3765
Asp Asp Gln Met Thr Leu Leu Gln Tyr Ser Trp Met Phe Leu Met Ala	885	890	895
TTT GCC TTG GGT TGG AGA TCA TAC AGA CAA TCA AGC GGA AAC CTG CTC			3813
Phe Ala Leu Gly Trp Arg Ser Tyr Arg Gln Ser Ser Gly Asn Leu Leu	900	905	910

TGC TTT GCT CCT GAT CTG ATT ATT AAT GAG CAG AGA ATG TCT CTA CCC Cys Phe Ala Pro Asp Leu Ile Ile Asn Glu Gln Arg Met Ser Leu Pro 915 920 925 930	3861
GGC ATG TAT GAC CAA TGT AAA CAC ATG CTG TTT GTC TCC TCT GAA TTA Gly Met Tyr Asp Gln Cys Lys His Met Leu Phe Val Ser Ser Glu Leu 935 940 945	3909
CAA AGA TTG CAG GTA TCC TAT GAA GAG TAT CTC TGT ATG AAA ACC TTA Gln Arg Leu Gln Val Ser Tyr Glu Glu Tyr Leu Cys Met Lys Thr Leu 950 955 960	3957
CTG CTT CTC TCC TCA GTT CCT AAG GAA GGT CTG AAG AGC CAA GAG TTA Leu Leu Leu Ser Ser Val Pro Lys Glu Gly Leu Lys Ser Gln Glu Leu 965 970 975	4005
TTT GAT GAG ATT CGA ATG ACT TAT ATC AAA GAG CTA GGA AAA GCC ATC Phe Asp Glu Ile Arg Met Thr Tyr Ile Lys Glu Leu Gly Lys Ala Ile 980 985 990	4053
GTC AAA AGG GAA GGG AAC TCC AGT CAG AAC TGG CAA CGG TTT TAC CAA Val Lys Arg Glu Gly Asn Ser Ser Gln Asn Trp Gln Arg Phe Tyr Gln 995 1000 1005 1010	4101
CTG ACA AAG CTT CTG GAC TCC ATG CAT GAG GTG GTT GAG AAT CTC CTT Leu Thr Lys Leu Leu Asp Ser Met His Glu Val Val Glu Asn Leu Leu 1015 1020 1025	4149
ACC TAC TGC TTC CAG ACA TTT TTG GAT AAG ACC ATG AGT ATT GAA TTC Thr Tyr Cys Phe Gln Thr Phe Leu Asp Lys Thr Met Ser Ile Glu Phe 1030 1035 1040	4197
CCA GAG ATG TTA GCT GAA ATC ATC ACT AAT CAG ATA CCA AAA TAT TCA Pro Glu Met Leu Ala Glu Ile Ile Thr Asn Gln Ile Pro Lys Tyr Ser 1045 1050 1055	4245
AAT GGA AAT ATC AAA AAG CTT CTG TTT CAT CAA AAA TGA CTGCCTTACT Asn Gly Asn Ile Lys Lys Leu Leu Phe His Gln Lys * 1060 1065 1070	4294
AAGAAAGGTT GCCTTAAAGA AAGTTGAATT TATAGTCTAG AGTCGACCCG GCGGCCGCT	4354
TCGAGCAGAC ATGATAAGAT ACATTGATGA GTTTGGACAA ACCACAATA GAATGCAGTG	4414
AAAAAATGC TTTATTTGTG AAATTTGTGA TGCTATTGCT TTATTTGTAA CCATTATAAG	4474
CTGCAATAAA CAAGTTAACA ACAACAATTG CATTCAATTTT ATGTTTCAGG TTCAGGGGGA	4534
GATGTGGGAG GTTTTTTAAA GCAAGTAAAA CCTCTACAAA TGTGGTAAAA TCGATAAGGA	4594
TCCGGGCTGG CGTAATAGCG AAGAGGCCCG CACCGATCGC CCTTCCCAAC AGTTGCGCAG	4654
CCTGAATGGC GAATGGACGC GCCCTGTAGC GGCGCATTAA GCGCGGCGGG TGTGGTGGTT	4714

ACGCGCAGCG	TGACCGCTAC	ACTTGCCAGC	GCCCTAGCGC	CCGCTCCTTT	CGCTTTCTTC	4774
CCTTCCTTTC	TCGCCACGTT	CGCCGGCTTT	CCCCGTCAAG	CTCTAAATCG	GGGGCTCCCT	4834
TTAGGGTTCC	GATTTAGAGC	TTTACGGCAC	CTCGACCGCA	AAAACTTGA	TTTGGGTGAT	4894
GGTTCACGTA	GTGGGCCATC	GCCCTGATAG	ACGGTTTTTC	GCCCTTTGAC	GTTGGAGTCC	4954
ACGTTCTTTA	ATAGTGGACT	CTTGTTCCAA	ACTGGAACAA	CACTCAACCC	TATCTCGGTC	5014
TATTCTTTTG	ATTTATAAGG	GATTTTGCCG	ATTTCGGCCT	ATTGGTTAAA	AAATGAGCTG	5074
ATTTAACAAA	TATTTAACGC	GAATTTTAAC	AAAATATTAA	CGTTTACAAT	TTCGCCTGAT	5134
GCGGTATTTT	CTCCTTACGC	ATCTGTGCGG	TATTTACACAC	CGCATATGGT	GCACTCTCAG	5194
TACAATCTGC	TCTGATGCCG	CATAGTTAAG	CCAGCCCCGA	CACCCGCCAA	CACCCGCTGA	5254
CGCGCCCTGA	CGGGCTTGTC	TGCTCCCGGC	ATCCGCTTAC	AGACAAGCTG	TGACCGTCTC	5314
CGGGAGCTGC	ATGTGTCAGA	GGTTTTCAAC	GTCATCACCG	AAACGCGCGA	GACGAAAGGG	5374
CCTCGTGATA	CGCCTATTTT	TATAGGTTAA	TGTCATGATA	ATAATGGTTT	CTTAGACGTC	5434
AGGTGGCACT	TTTCGGGGAA	ATGTGCGCGG	AACCCCTATT	TGTTTATTTT	TCTAAATACA	5494
TTCAAATATG	TATCCGCTCA	TGAGACAATA	ACCCTGATAA	ATGCTTCAAT	AATATTGAAA	5554
AAGGAAGAGT	ATGAGTATTC	AACATTTCCG	TGTCGCCCTT	ATTCCCTTTT	TTGCGGCATT	5614
TTGCCTTCCT	GTTTTTGCTC	ACCCAGAAAC	GCTGGTGAAA	GTAAAAGATG	CTGAAGATCA	5674
GTTGGGTGCA	CGAGTGGGTT	ACATCGAACT	GGATCTCAAC	AGCGGTAAGA	TCCTTGAGAG	5734
TTTTCGCCCC	GAAGAACGTT	TTCCAATGAT	GAGCACTTTT	AAAGTTCTGC	TATGTGGCGC	5794
GGTATTATCC	CGTATTGACG	CCGGGCAAGA	GCAACTCGGT	CGCCGCATAC	ACTATTCTCA	5854
GAATGACTTG	GTTGAGTACT	CACCAGTCAC	AGAAAAGCAT	CTTACGGATG	GCATGACAGT	5914
AAGAGAATTA	TGCAGTGCTG	CCATAACCAT	GAGTGATAAC	ACTGCGGCCA	ACTTACTTCT	5974
GACAACGATC	GGAGGACCGA	AGGAGCTAAC	CGCTTTTTTG	CACAACATGG	GGGATCATGT	6034
AACTCGCCTT	GATCGTTGGG	AACCGGAGCT	GAATGAAGCC	ATACCAAACG	ACGAGCGTGA	6094
CACCACGATG	CCTGTAGCAA	TGGCAACAAC	GTTGCGCAAA	CTATTAAGTG	GCGAACTACT	6154
TACTCTAGCT	TCCCGGCAAC	AATTAATAGA	CTGGATGGAG	GCGGATAAAG	TTGCAGGACC	6214
ACTTCTGCGC	TCGGCCCTTC	CGGCTGGCTG	GTTTATTGCT	GATAAATCTG	GAGCCGGTGA	6274
GCGTGGGTCT	CGCGGTATCA	TTGCAGCACT	GGGGCCAGAT	GGTAAGCCCT	CCCGTATCGT	6334
AGTTATCTAC	ACGACGGGGA	GTCAGGCAAC	TATGGATGAA	CGAAATAGAC	AGATCGCTGA	6394
GATAGGTGCC	TCACTGATTA	AGCATTTGTA	ACTGTCAGAC	CAAGTTTACT	CATATATACT	6454

TTAGATTGAT TTAAAACTTC ATTTTAAATT TAAAAGGATC TAGGTGAAGA TCCTTTTTGA	6514
TAATCTCATG ACCAAAATCC CTTAACGTGA GTTTTCGTTC CACTGAGCGT CAGACCCCGT	6574
AGAAAAGATC AAAGGATCTT CTTGAGATCC TTTTTTCTG CGCGTAATCT GCTGCTTGCA	6634
AACAAAAAAA CCACCGCTAC CAGCGGTGGT TTGTTTGCCG GATCAAGAGC TACCAACTCT	6694
TTTTCCGAAG GTAAGTGGCT TCAGCAGAGC GCAGATACCA AATACTGTCC TTCTAGTGTA	6754
GCCGTAGTTA GGCCACCACT TCAAGAACTC TGTAGCACCG CCTACATACC TCGCTCTGCT	6814
AATCCTGTTA CCAGTGGCTG CTGCCAGTGG CGATAAGTCG TGTCTTACCG GGTTGGACTC	6874
AAGACGATAG TTACCGGATA AGGCGCAGCG GTCGGGCTGA ACGGGGGGTT CGTGCACACA	6934
GCCCAGCTTG GAGCGAACGA CCTACACCGA ACTGAGATAC CTACAGCGTG AGCTATGAGA	6994
AAGCGCCACG CTTCCCGAAG GGAGAAAGGC GGACAGGTAT CCGGTAAGCG GCAGGGTCGG	7054
AACAGGAGAG CGCACGAGGG AGCTTCCAGG GGGAAACGCC TGGTATCTTT ATAGTCCTGT	7114
CGGGTTTCGC CACCTCTGAC TTGAGCGTCG ATTTTGTGA TGCTCGTCAG GGGGGCGGAG	7174
CCTATGGAAG AACGCCAGCA ACGCGGCCTT TTTACGGTTC CTGGCCTTTT GCTGGCCTTT	7234
TGCTCACATG GCTCGACAGA TCT	7257

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1071 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Ala	His	His	His	His	His	His	Gly	Tyr	Pro	Tyr	Asp	Val	Pro	Asp
1				5					10					15	
Tyr	Ala	Gln	Ser	Ser	Ala	Met	Ser	Lys	Gly	Glu	Glu	Leu	Phe	Thr	Gly
			20					25					30		
Val	Val	Pro	Ile	Leu	Val	Glu	Leu	Asp	Gly	Asp	Val	Asn	Gly	His	Lys
		35					40					45			
Phe	Ser	Val	Ser	Gly	Glu	Gly	Glu	Gly	Asp	Ala	Thr	Tyr	Gly	Lys	Leu
	50					55					60				
Thr	Leu	Lys	Phe	Ile	Cys	Thr	Thr	Gly	Lys	Leu	Pro	Val	Pro	Trp	Pro
65					70				75					80	
Thr	Leu	Val	Thr	Thr	Phe	Thr	Tyr	Gly	Val	Gln	Cys	Phe	Ser	Arg	Tyr

85										90					95				
Pro	Asp	His	Met	Lys	Gln	His	Asp	Phe	Phe	Lys	Ser	Ala	Met	Pro	Glu				
			100					105					110						
Gly	Tyr	Val	Gln	Glu	Arg	Thr	Ile	Phe	Phe	Lys	Asp	Asp	Gly	Asn	Tyr				
		115					120					125							
Lys	Thr	Arg	Ala	Glu	Val	Lys	Phe	Glu	Gly	Asp	Thr	Leu	Val	Asn	Arg				
		130				135					140								
Ile	Glu	Leu	Lys	Gly	Ile	Asp	Phe	Lys	Glu	Asp	Gly	Asn	Ile	Leu	Gly				
145					150					155					160				
His	Lys	Leu	Glu	Tyr	Asn	Tyr	Asn	Ser	His	Asn	Val	Tyr	Ile	Met	Ala				
				165					170					175					
Asp	Lys	Gln	Lys	Asn	Gly	Ile	Lys	Val	Asn	Phe	Lys	Ile	Arg	His	Asn				
			180					185					190						
Ile	Glu	Asp	Gly	Ser	Val	Gln	Leu	Ala	Asp	His	Tyr	Gln	Gln	Asn	Thr				
		195				200						205							
Pro	Ile	Gly	Asp	Gly	Pro	Val	Leu	Leu	Pro	Asp	Asn	His	Tyr	Leu	Ser				
		210				215					220								
Thr	Gln	Ser	Ala	Leu	Ser	Lys	Asp	Pro	Asn	Glu	Lys	Arg	Asp	His	Met				
225					230					235					240				
Val	Leu	Leu	Glu	Phe	Val	Thr	Ala	Ala	Gly	Ile	Thr	His	Gly	Met	Asp				
				245					250					255					
Glu	Leu	Tyr	Lys	Gly	Ala	Gly	Ala	Gly	Ala	Gly	Ala	Gly	Ala	Ile	Ser				
			260					265					270						
Ala	Leu	Ile	Leu	Asp	Ser	Lys	Glu	Ser	Leu	Ala	Pro	Pro	Gly	Arg	Asp				
		275					280					285							
Glu	Val	Pro	Gly	Ser	Leu	Leu	Gly	Gln	Gly	Arg	Gly	Ser	Val	Met	Asp				
		290				295					300								
Phe	Tyr	Lys	Ser	Leu	Arg	Gly	Gly	Ala	Thr	Val	Lys	Val	Ser	Ala	Ser				
305					310					315					320				
Ser	Pro	Ser	Val	Ala	Ala	Ala	Ser	Gln	Ala	Asp	Ser	Lys	Gln	Gln	Arg				
				325					330					335					
Ile	Leu	Leu	Asp	Phe	Ser	Lys	Gly	Ser	Thr	Ser	Asn	Val	Gln	Gln	Arg				
			340					345					350						
Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln				
		355					360					365							
Gln	Gln	Gln	Pro	Gly	Leu	Ser	Lys	Ala	Val	Ser	Leu	Ser	Met	Gly	Leu				
		370				375					380								
Tyr	Met	Gly	Glu	Thr	Glu	Thr	Lys	Val	Met	Gly	Asn	Asp	Leu	Gly	Tyr				

385		390		395		400
Pro Gln Gln Gly	Gln Leu Gly Leu Ser Ser Gly Glu Thr Asp Phe Arg					
	405		410		415	
Leu Leu Glu Glu Ser Ile Ala Asn Leu Asn Arg Ser Thr Ser Val Pro						
	420		425		430	
Glu Asn Pro Lys Ser Ser Thr Ser Ala Thr Gly Cys Ala Thr Pro Thr						
	435		440		445	
Glu Lys Glu Phe Pro Lys Thr His Ser Asp Ala Ser Ser Glu Gln Gln						
	450		455		460	
Asn Arg Lys Ser Gln Thr Gly Thr Asn Gly Gly Ser Val Lys Leu Tyr						
	465		470		475	480
Pro Thr Asp Gln Ser Thr Phe Asp Leu Leu Lys Asp Leu Glu Phe Ser						
	485		490		495	
Ala Gly Ser Pro Ser Lys Asp Thr Asn Glu Ser Pro Trp Arg Ser Asp						
	500		505		510	
Leu Leu Ile Asp Glu Asn Leu Leu Ser Pro Leu Ala Gly Glu Asp Asp						
	515		520		525	
Pro Phe Leu Leu Glu Gly Asn Thr Asn Glu Asp Cys Lys Pro Leu Ile						
	530		535		540	
Leu Pro Asp Thr Lys Pro Lys Ile Lys Asp Thr Gly Asp Thr Ile Leu						
	545		550		555	560
Ser Ser Pro Ser Ser Val Ala Leu Pro Gln Val Lys Thr Glu Lys Asp						
	565		570		575	
Asp Phe Ile Glu Leu Cys Thr Pro Gly Val Ile Lys Gln Glu Lys Leu						
	580		585		590	
Gly Pro Val Tyr Cys Gln Ala Ser Phe Ser Gly Thr Asn Ile Ile Gly						
	595		600		605	
Asn Lys Met Ser Ala Ile Ser Val His Gly Val Ser Thr Ser Gly Gly						
	610		615		620	
Gln Met Tyr His Tyr Asp Met Asn Thr Ala Ser Leu Ser Gln Gln Gln						
	625		630		635	640
Asp Gln Lys Pro Val Phe Asn Val Ile Pro Pro Ile Pro Val Gly Ser						
	645		650		655	
Glu Asn Trp Asn Arg Cys Gln Gly Ser Gly Glu Asp Ser Leu Thr Ser						
	660		665		670	
Leu Gly Ala Leu Asn Phe Pro Gly Arg Ser Val Phe Ser Asn Gly Tyr						
	675		680		685	

Ser Ser Pro Gly Met Arg Pro Asp Val Ser Ser Pro Pro Ser Ser Ser
 690 695 700

Ser Ala Ala Thr Gly Pro Pro Pro Lys Leu Cys Leu Val Cys Ser Asp
 705 710 715 720

Glu Ala Ser Gly Cys His Tyr Gly Val Leu Thr Cys Gly Ser Cys Lys
 725 730 735

Val Phe Phe Lys Arg Ala Val Glu Gly Gln His Asn Tyr Leu Cys Ala
 740 745 750

Gly Arg Asn Asp Cys Ile Ile Asp Lys Ile Arg Arg Lys Asn Cys Pro
 755 760 765

Ala Cys Arg Tyr Arg Lys Cys Leu Gln Ala Gly Met Asn Leu Glu Ala
 770 775 780

Arg Lys Thr Lys Lys Lys Ile Lys Gly Ile Gln Gln Ala Thr Ala Gly
 785 790 795 800

Val Ser Gln Asp Thr Ser Glu Asn Pro Asn Lys Thr Ile Val Pro Ala
 805 810 815

Ala Leu Pro Gln Leu Thr Pro Thr Leu Val Ser Leu Leu Glu Val Ile
 820 825 830

Glu Pro Glu Val Leu Tyr Ala Gly Tyr Asp Ser Ser Val Pro Asp Ser
 835 840 845

Ala Trp Arg Ile Met Thr Thr Leu Asn Met Leu Gly Gly Arg Gln Val
 850 855 860

Ile Ala Ala Val Lys Trp Ala Lys Ala Ile Leu Gly Leu Arg Asn Leu
 865 870 875 880

His Leu Asp Asp Gln Met Thr Leu Leu Gln Tyr Ser Trp Met Phe Leu
 885 890 895

Met Ala Phe Ala Leu Gly Trp Arg Ser Tyr Arg Gln Ser Ser Gly Asn
 900 905 910

Leu Leu Cys Phe Ala Pro Asp Leu Ile Ile Asn Glu Gln Arg Met Ser
 915 920 925

Leu Pro Gly Met Tyr Asp Gln Cys Lys His Met Leu Phe Val Ser Ser
 930 935 940

Glu Leu Gln Arg Leu Gln Val Ser Tyr Glu Glu Tyr Leu Cys Met Lys
 945 950 955 960

Thr Leu Leu Leu Leu Ser Ser Val Pro Lys Glu Gly Leu Lys Ser Gln
 965 970 975

Glu Leu Phe Asp Glu Ile Arg Met Thr Tyr Ile Lys Glu Leu Gly Lys
 980 985 990

Ala Ile Val Lys Arg Glu Gly Asn Ser Ser Gln Asn Trp Gln Arg Phe
995 1000 1005

Tyr Gln Leu Thr Lys Leu Leu Asp Ser Met His Glu Val Val Glu Asn
1010 1015 1020

Leu Leu Thr Tyr Cys Phe Gln Thr Phe Leu Asp Lys Thr Met Ser Ile
1025 1030 1035 1040

Glu Phe Pro Glu Met Leu Ala Glu Ile Ile Thr Asn Gln Ile Pro Lys
1045 1050 1055

Tyr Ser Asn Gly Asn Ile Lys Lys Leu Leu Phe His Gln Lys *
1060 1065 1070

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 46 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GCGCGCTGAT CAGAATTCCT TTTAGGAATT CTGATCAGCG CGCTGA

46

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

AGAACANNNT GTTCT

15

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AGGTCANNNT GACCT

15

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 60 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

TCGAGCGCGC AAGAACACAG TGTTC TGACG ACACGAAGAA CAGGATGTTT TCGTACAGTG

60

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 60 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

TCGACACTGT ACGAGAACAT CCTGTTCTTC GTGTCGTCAG AACACTGTGT TCTTGCGCGC

60

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 60 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

TCGAGCGCGC AAGGTCACAG TGACCTGACG ACACGAAGGT CAGGATGACC TCGTACAGTG

60

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

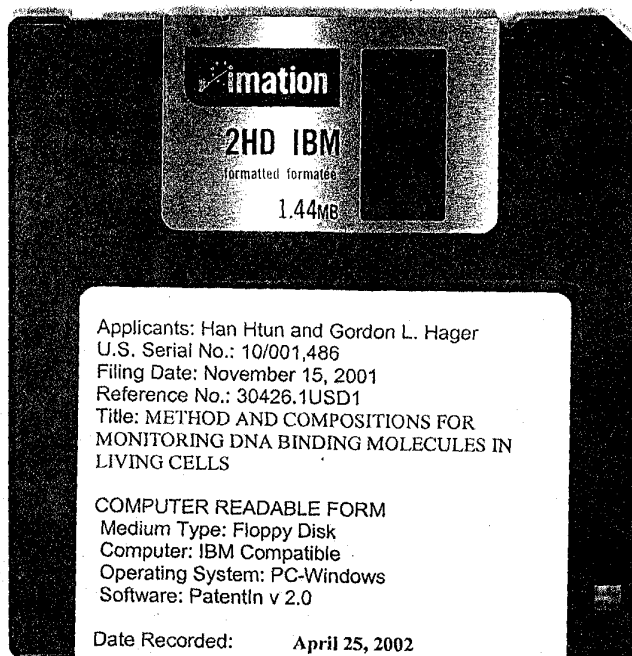
- (A) LENGTH: 60 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TCGACACTGT ACGAGGTCAT CCTGACCTTC GTGTCGTCAG GTCAGTGTGA CCTTGCGCGC

60



Applicants: Han Htun and Gordon L. Hager
U.S. Serial No.: 10/001,486
Filing Date: November 15, 2001
Reference No.: 30426.1USD1
Title: METHOD AND COMPOSITIONS FOR
MONITORING DNA BINDING MOLECULES IN
LIVING CELLS

COMPUTER READABLE FORM
Medium Type: Floppy Disk
Computer: IBM Compatible
Operating System: PC-Windows
Software: PatentIn v 2.0

Date Recorded: April 25, 2002